运动想象脑电信号的跨域特征学习方法 *

韦泓妤[a]，陈黎飞[a,b,c†]，罗天健[a,b]

(福建师范大学 a. 计算机与网络空间安全学院; b. 数字福建环境监测物联网实验室; c. 福建省应用数学中心, 福州 350108)

摘　要：运动想象脑电信号采集成本高且个体差异大，跨个体域构建脑电信号模式识别模型属于典型的小样本跨域学习任务。针对该任务，提出了一种运动想象脑电信号的跨域特征学习方法。该方法首先选择最优度量方法对齐协方差并提取共同空间模式特征；其次，在该特征基础上采用领域自适应方法学习目标域的最优跨域特征。为验证所提出方法的可行性与有效性，采用经典模型识别跨域特征，在两个公开的数据集上进行对比实验。实验结果表明，通过所提出方法学习到的跨域特征，在运动想象模式识别中，明显优于现有方法学习到的特征。此外，还详细对比了跨域特征学习方法的各项参数设置、性能及效率。

关键词：运动想象；脑电信号；跨域特征学习；领域自适应；协方差对齐

中图分类号：TP391　　doi: 10.19734/j.issn.1001-3695.2022.01.0016

Cross-domain feature learning method for motor imagery EEG signals

Wei Hongyu[a], Chen Lifei[a, b, c†], Luo Tianjian[a, b]

(a. *College of Computer & Cyber Security*, b. *Digital Fujian Environmental Monitoring Internet of Things Laboratory*, c. *Fujian Applied Mathematics Center*, *Fujian Normal University* ,*Fuzhou* 350108, *China*)

Abstract: Motor imagery EEG signals requires a high cost for recording, and there is a large difference for individual's signals. Cross-subject motor imagery EEG signals recognition task belongs to a typical cross-domain learning problem with a small samples set. To solve this problem, this paper proposed a cross-domain feature learning method for motor imagery EEG signals to improve the recognition performance. The proposed method first selected the optimal measurement to align the covariance of EEG signals, and then extracted common spatial patterns (CSP) from the aligned EEG trials. Second, based on the CSP features, an optimal domain adaptation algorithm was selected to learn the optimal cross-domain features for the target domain. To validate the feasibility and effectiveness of the learned cross-domain features, a classical model was adopted to recognize the learned cross-domain features, and the comparative experiments were conducted on two public datasets. Experimental results show that the learned cross-domain features are obviously better than state-of-the-arts methods in recognition performance. In addition, this paper also compared the parameters setting, performance and efficiency for the proposed method.

Key words: motor imagery; EEG signal; cross-domain feature learning; domain adaptation; covariance alignment

0　引言

脑机接口(BCI)[1,2]技术绕过肌肉系统，允许大脑与外部设备之间建立直接的通路，实现神经系统和外部设备间的直接信息通信。BCI 技术分为三个主要步骤：收集大脑信号，分析并解码大脑信号，再根据解码结果向外部设备输出控制命令。头皮脑电(EEG)是 BCI 最常用的脑电信号类型之一，该信号采用非侵入方式记录大脑活动的电生理现象。基于 EEG 的 BCI 主要包括三种方式：运动想象(MI)，稳态视觉电位(SSVEP)、事件相关电位(ERP)。其中，MI 属于主动 BCI 模式，因此基于 EEG 的 MI-BCI 是最常使用的 BCI 范式，也是本文的研究对象[3]。

在 MI-BCI 中，用户根据提示想象自己身体部位的运动，此时会在对应的大脑皮层区域引起感知运动节律，产生大脑相关皮层区域的脑节律调制，从而在 EEG 信号中编码不同的运动想象规律和特性。因此，不同运动动作的想象可以产生不同的空间局部化解码，从而识别出不同的运动感知节律模式，最终可根据模式识别结果的映射关系控制外部设备。早期 MI-BCI 已经应用至医学领域，用于帮助严重瘫痪的患者在没有肌肉系统参与的情况下控制轮椅[4]。近年来，MI-BCI 已经帮助中风患者做康复训练以改善其运动功能[5,6]，并且 MI-BCI 的娱乐性应用已经扩展至健康人群。然而，EEG 信号记录人脑头皮层的微弱信号变化，易受噪声影响，常用的特征提取方法为共同空间模式(CSP)。CSP 是应用非常广泛的一种空域滤波特征提取算法[7,8]，在运动想象任务的特征提取与分类中有不错的效果，该方法能够从多通道的 EEG 信号中提取出每个类别的最优空间分布成分。在 CSP 特征基础上，一般可采用线性判别分析(LDA)或支持向量机(SVM)等模型完成运动想象任务的模式识别。

实际上，由于 EEG 信号采集的成本较高，通常单个被试者的 EEG 样本数量较少，并且每个被试者的 EEG 信号分布存在个体差异性，无法将跨被试的全体 EEG 信号直接用于构建高性能模式识别模型。因此，一些研究者希望通过迁移学习的方法降低EEG信号特征分布的个体差异性[9]，并在全体跨被试者的 EEG 信号上构建高性能模式识别模型。迁移学习中最常见的任务为领域自适应，经典的领域自适应方法包括 Yang 等提出的迁移成分分析(TCA)[10]、Long 等提出的联合分布适配方法(JDA)[11]、以及考虑稀疏特征选择的迁

收稿日期：2022-01-15；**修回日期：**2022-03-11　　**基金项目：**国家自然科学基金资助项目(U1805263)；国家自然科学基金青年资助项目(62106049)

作者简介：韦泓妤(1999-)，女，广西桂平人，硕士研究生，主要研究方向为脑机接口；陈黎飞(1972-)，男(通信作者)，福建长乐人，教授，博士，主要研究方向为统计机器学习、数据挖掘、模式识别(clfei@fjnu.edu.cn)；罗天健(1990-)，男，湖北黄冈人，讲师，博士，主要研究方向为脑机接口与模式识别.

移联合匹配改进方法(TJM)[12]，还包括 Wang 等提出的动态分布自适应(BDA)[13]。在解决领域自适应的方法中，其基本目标是降低源域和目标域的特征分布差异，以保证源域和目标域的特征分布能共同用于训练分类器，这些方法可拓展至 MI-BCI 的 CSP 特征之上，用以提升模式识别的性能[14,15]。

在基于 MI-BCI 的迁移学习中，一些研究者[16]结合核共同空间模式(KCSP)[17]和迁移核学习(TKL)[18]提出了迁移核共同空间模式方法(TKCSP)，并应用于跨被试者的脑电信号处理中。Jayaram[19]等研究者提出了一种不需要目标域标记数据的跨被试者迁移学习的多任务学习框架，并证明该方法对一名肌肉萎缩患者的迁移效果不错。Azab[20]等研究者对上述方法做了进一步的改进，提出 MI 分类中对跨被试者利用 Kullback-Leibler 散度进行加权迁移学习。在集成迁移学习中，Hossain[21]等研究者提出了多类 MI 分类中跨被试迁移的集成学习方法。

在数据对齐方面，Zanini[22]等研究者提出了一种黎曼距离对齐方法(RA)，并采用改进的 MDRM[23]分类器验证了该方法的有效性。Yair[24]等研究者提出了一种在 SPD 矩阵的锥流形上，通过平行输运解析表达式的领域自适应方法，并证明该方法在跨被试者运动想象分类、睡眠阶段分类和心算识别方面有不错的效果。He 和 Wu[25]在 RA 的基础上做了改进，提出了一种欧式距离对齐方法(EA)，该方法在欧氏空间上将不同被试者的 EEG 信号协方差矩阵对齐到单位矩阵上，比 RA 方法更方便、高效。此外，Rodrigues[26]等研究者提出了黎曼 Procrustes 分析(RPA)方法，用于对齐跨被试者的 EEG 信号，该方法在跨被试 MI、ERP 和 SSVEP 分类中有着不错的效果。

在流形迁移学习中，Zhang 和 Wu[27]在黎曼流形协方差矩阵对齐基础上，提出了一种在切空间上的流形知识迁移方法(MEKT)。Singh[28]等研究者提出了一种样本选择的迁移学习方法，该方法从少量的目标域样本中估计 MDRM 分类器使用的样本协方差矩阵。另外，基于深度学习的 MI-BCI 迁移学习[29]也获得了不错的效果。Wu 等研究者[30]提出了一种应用于 MI-BCI 分类的并行多尺度滤波器(CNN)，提升了分类结果在迁移上的性能。

运动想象 EEG 信号具有非平稳特性且易受噪声影响，获取个体域记录的信号代价高且特征分布差异大，跨个体域构建脑电信号模式识别模型属于典型的小样本跨域学习任务。从现有研究的综述中可以看出，目前的研究集中在样本协方差对齐、流形特征自适应或模式识别模型的迁移[31]。为了在个体域之间学习稳定、可靠的跨域特征，本文提出了一种运动想象脑电信号的跨域特征学习方法，用于提升跨个体域脑电信号模式识别性能。首先，该方法选择最优的协方差矩阵距离度量方法，在个体域之间将样本集协方差对齐至单位矩阵。其次，在对齐后的样本集上提取CSP特征，并引入TCA、JDA、BDA 和 TJM 四种不同的领域自适应方法，在待识别的目标域上学习最佳的跨域特征。本文在两个公开数据集上验证了所提出跨域特征学习方法的可行性与有效性，并且从参数设置、模式识别性能和效率上进行了详细对比实验分析。

现有面向运动想象脑机接口的跨域特征学习方法仅针对样本对齐角度或 CSP 特征跨域自适应角度，跨域的特征学习能力不够，导致运动想象模式识别性能不足。本文同时针对样本角度和 CSP 特征角度构建跨域特征学习方法，为个体域脑电信号样本集选择最优的样本对齐度量方法，以及 CSP 特征跨域自适应方法，并在两个公开的脑电数据集上验证了所提出方法的模式识别性能及效率。

1　最优协方差矩阵样本对齐

在跨个体域的 EEG 信号样本对齐中，本文选用欧式距离对齐方法(EA)[25]。由于该方法需将样本对齐至协方差中心，因此本文将协方差中心计算扩展至不同距离度量方法，并为不同个体选择最优的协方差中心度量方法。

1.1　协方差矩阵中心对齐方法

协方差矩阵中心可作为多维时间序列分布的度量方式。假设 $\boldsymbol{X}_i$ 是被试者进行运动想象任务获得的第 i 个样本，每个被试者拥有 n 个 EEG 样本，则个体域的 EEG 信号样本协方差矩阵中心 $\bar{\boldsymbol{R}}$ 为：$\bar{\boldsymbol{R}}=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i\boldsymbol{X}_i^T$。该中心采用算法平均值表达，在经典的 EA 算法中使用欧式距离度量。随后，根据协方差矩阵中心，对每个 EEG 样本进行对齐操作：

$$\tilde{\boldsymbol{X}}_i=\bar{\boldsymbol{R}}^{-\frac{1}{2}}\boldsymbol{X}_i \tag{1}$$

经过 n 次对齐后，所有被试者的 EEG 样本的平均协方差矩阵都等于单位矩阵，因此来自不同被试者的协方差矩阵分布会更为相似。

1.2　最优协方差矩阵中心距离度量方法

由于不同任务、不同被试者对 EEG 信号样本的分布都有影响，在实际个体域的对齐中，需根据每个个体域的样本分布情况，灵活选择合适的协方差中心度量方法。除了欧式距离以外，本文将采用如下 7 种均值度量方法：

(1)黎曼均值：

$$\bar{\boldsymbol{X}}=argmin\left(\sum_i\delta_R(\boldsymbol{X},\boldsymbol{X}_i)^2\right) \tag{2}$$

其中：δ_R 表示黎曼空间上两个协方差矩阵的距离。采用梯度下降法，在黎曼空间上进行迭代的均值求解。

(2)对数欧式均值：

$$\bar{\boldsymbol{X}}=\exp\left(\frac{1}{n}\sum_i\log\boldsymbol{X}_i\right) \tag{3}$$

(3)对数 det 均值：

$$\bar{\boldsymbol{X}}=\left(\sum_i(0.5X+0.5\boldsymbol{X}_i)^{-1}\right)^{-1} \tag{4}$$

(4) Wasserstein 均值：

$$\boldsymbol{K}=\left(\sum_i(\boldsymbol{K}\boldsymbol{X}_i\boldsymbol{K})^{1/2}\right)^{1/2} \tag{5}$$

其中：$\boldsymbol{K}=\bar{\boldsymbol{X}}^{-1/2}$。根据 Wasserstein 度量方式，初始化迭代次数，不断更新 $\boldsymbol{K}$，直到迭代次数停止。

(5) AJD 对数均值：

$$\bar{\boldsymbol{X}}=\boldsymbol{A}^T exp\left(\sum_i\log\boldsymbol{B}^T\boldsymbol{X}_i\boldsymbol{B}\right)\boldsymbol{A} \tag{6}$$

其中：矩阵 $\boldsymbol{B}$ 是基于 Pham 算法的近似联合对角化，$\boldsymbol{A}$ 是 $\boldsymbol{B}$ 的逆矩阵。

(6)调和均值：

$$\bar{\boldsymbol{X}}=\left(\frac{1}{n}\sum_i\boldsymbol{X}_i{}^{-1}\right)^{-1} \tag{7}$$

(7) KL 散度均值，计算方式为

$$\bar{\boldsymbol{X}}=(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i)^{1/2}\left[(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i)^{-1/2}\left(\frac{1}{n}\sum_i\boldsymbol{X}_i{}^{-1}\right)^{-1}\right]^{1/2}(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i)^{1/2} \tag{8}$$

实际上，该均值是算术平均值和调和平均值之间的几何平均值。

实际实验中，本文将根据不同个体域 EEG 样本分布情况，采用 trial-and-error 的方式选择最优的度量方式，计算协方差矩阵中心算术平均值，并采用式(1)对齐不同个体域的 EEG 信号样本集。

2　EEG 信号的跨域 CSP 特征学习

在迁移学习方法中，领域自适应是常见的基于核空间的方法。其本质是最小化源域特征分布和目标域特征分布的距离，最常使用衡量源域特征分布和目标域特征分布的距离为

MMD [32]。在 EEG 信号的跨域 CSP 特征学习中，本文的目标是在多个源域和目标域 CSP 特征最小化 *MMD* 距离。

2.1　迁移成分分析方法

迁移成分分析方法(TCA)为基于边缘分布的自适应[10]方法。假设用源域和目标域的边缘分布采用 $P(X_s)$ 和 $P(X_t)$ 表示，TCA 方法的目的是减小 $P(X_s)$ 和 $P(X_t)$ 之间的差异。TCA 的目的是最小化 *MMD* 距离，其优化目标为

$$\min_{w} tr(\boldsymbol{A}^T\boldsymbol{XLX}^T\boldsymbol{A})+\mu\|\boldsymbol{A}\|_F^2 \quad s.t.\ \ \boldsymbol{A}^T\boldsymbol{XHX}^T\boldsymbol{A}=\boldsymbol{I} \tag{9}$$

其中，$\boldsymbol{A}$ 是最小化 MMD 的映射矩阵，$\boldsymbol{H}$ 是中心矩阵，$\boldsymbol{H}=\boldsymbol{I}_{n_1+n_2}-1/(n_1+n_2)11^T$，$\boldsymbol{L}$ 即为 *MMD* 距离。采用拉格朗日乘子法即可求得 $\boldsymbol{A}$ 的最优解为 $(\boldsymbol{XLX}+\mu\boldsymbol{I})^{-1}\boldsymbol{XHX}$ 最大的 m 个特征值对应的特征向量。最终，通过映射 $\boldsymbol{A}$ 变换源域和目标域特征。

2.2　联合分布自适应

联合分布自适应方法 (JDA)[11]的目标是减小源域和目标域的联合概率分布的距离。用 $P(X_s)$ 和 $P(X_t)$ 之间的距离、以及 $P(Y_s|X_s)$ 和 $P(Y_t|X_t)$ 之间的距离来近似表达。由于目标域没有标记数据，所以无法直接求解 $P(Y_t|X_t)$。JDA 方法通过迭代的方式在源域特征和样本上训练简单的分类器，并在目标域特征上预测伪标签 y_t'。形式地，JDA 的优化目标为

$$\min_{A}\sum_{c=0}^{C} tr(\boldsymbol{A}^T\boldsymbol{XM}_c\boldsymbol{X}^T\boldsymbol{A})+\lambda\|\boldsymbol{A}\|_F^2 \quad s.t.\ \ \boldsymbol{A}^T\boldsymbol{XHX}^T\boldsymbol{A}=\boldsymbol{I} \tag{10}$$

其中：$\boldsymbol{A}^T\boldsymbol{XHX}^T\boldsymbol{A}=\boldsymbol{I}$，$\boldsymbol{H}$ 是中心矩阵，$\boldsymbol{I}$ 是单位矩阵，$\boldsymbol{M}_c$ 表示由源域样本和标签，以及目标上样本和预测伪标签计算的 *MMD*。通过同样计算方法，获得映射矩阵 $\boldsymbol{A}$，用于变换源域和目标域特征。

2.3　动态分布自适应

JDA 方法将边缘分布和概率分布的重要性视作同等，然而，由于不同 BCI 任务和不同人群，其重要性并不始终相等。针对 JDA 的不足之处，动态分布自适应方法(BDA)设置平衡因子 $\mu\in[0,1]$ 动态调整边缘分布和条件分布的重要性[13]。形式地，BDA 的优化目标为

$$\min_{A} tr\left(\boldsymbol{A}^T\boldsymbol{X}\left((1-\mu)\boldsymbol{M}_0+\mu\sum_{c=1}^{C}\boldsymbol{M}_c\right)\boldsymbol{X}^T\boldsymbol{A}\right)+\lambda\|\boldsymbol{A}\|_F^2 \quad s.t.\ \ \boldsymbol{A}^T\boldsymbol{XHX}^T\boldsymbol{A}=\boldsymbol{I},0\le\mu<1 \tag{11}$$

其中：BDA 方法采用 A-distance 估计。需要指出的是，BDA 方法每轮迭代后，利用伪标签和 A-distance 计算新的 μ 值。同样地，获得的映射矩阵用于变换源域和目标域特征。

2.4　迁移联合匹配

无论是 JDA 还是 BDA,都将源域中所有的特征整体用于计算 *MMD*，从而获得最优特征映射矩阵 $\boldsymbol{A}$。实际上，源域中只有一些有迁移特征意义的特征，对计算 *MMD* 能够起到决定作用，而另一些是存在负迁移的特征，则会对 *MMD* 计算造成反作用。因此，TJM 方法[12]在计算放射矩阵时，将源域的放射矩阵添加稀疏正则化 $\|\boldsymbol{A}_s\|_{2,1}$，而保证目标域的放射矩阵为二范数不变 $\|\boldsymbol{A}_t\|_F^2$。通过系数正则化的选择，即可在优化过程中从源域中选择更具代表性的特征，用于计算 *MMD*,从而提升特征迁移的性能[33]。因此，在 TCA 基础上 TJM 的定义可写为

$$\min_{A} tr(\boldsymbol{A}^T\boldsymbol{KMK}^T\boldsymbol{A})+\lambda\left(\|\boldsymbol{A}_s\|_{2,1}+\|\boldsymbol{A}_t\|_F^2\right) \quad s.t.\ \ \boldsymbol{A}^T\boldsymbol{KHK}^T\boldsymbol{A}=\boldsymbol{I} \tag{12}$$

其中：λ 为正则项的系数。由于上述优化目标具有非凸的特征，因此需要通过迭代的方式求解。最终，由 $\boldsymbol{A}=[\boldsymbol{A}_s,\boldsymbol{A}_t]$ 构建稀疏映射矩阵，在变换源域和目标域特征的同时保证稀疏性，提取最具代表性特征　。

3　运动想象 EEG 信号跨域特征学习方法

3.1　共同空间模式

共同空间模式(CSP)[34]针对二分类任务，从多通道 EEG 信号中提取可区分特征。假设 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 分别为两类不同运动想象任务的 EEG 信号，首先，计算各域对齐后 EEG 信号样本集的平均协方差矩阵 $\boldsymbol{C}_i\ (i=1,2)$；其次，寻找一个最优的空间滤波器 $\boldsymbol{w}$,使得类别 1 的方差最大，类别 2 的方差最小，其优化目标为[34]：

$$\boldsymbol{C}_1\boldsymbol{w}=\lambda(\boldsymbol{C}_1+\boldsymbol{C}_2)\boldsymbol{w} \tag{13}$$

其中：λ 和 $\boldsymbol{w}$ 分别为广义特征值和特征向量，然后利用 $\boldsymbol{Z}_{M\times N}=\boldsymbol{W}_{M\times M}\boldsymbol{X}_{M\times N}$ 将 EEG 信号进行空域滤波获得特征向量 $\boldsymbol{Z}$。随后，选择特征向量 $\boldsymbol{Z}$ 的前 m 行和后 m 行 $(2\mathrm{m}<M)$，形成的特征向量作为分离后的空域滤波特征。对于 EEG 信号 $\boldsymbol{X}_i$，CSP 算法最终提取到的特征 $\boldsymbol{f}_i$ 为：$\boldsymbol{f}_i=log(\boldsymbol{ZZ}^T)$。

3.2　线性判别分析

线性判别分析(LDA)[35]是 MI-BCI 中常用的分类器，对提取到的 CSP 特征分类后，完成运动想象任务的模式识别。给定特征集合 $D=\{(x_i,y_i)\}_{i=1}^{m},y_i\in\{0,1\}$，$\boldsymbol{X}_i,\boldsymbol{\mu}_i,\Sigma_i$ 分别表示第 $i\in\{1,2\}$ 类运动想象 EEG 信号提取到的跨域特征，均值向量和协方差矩阵。首先，LDA 分别计算为“类内散度矩阵” $\boldsymbol{S}_b$ 和“类间散度矩阵” $\boldsymbol{S}_w$，并保证 $\boldsymbol{S}_b$ 足够大且 $\boldsymbol{S}_w$ 足够小，其优化目标为

$$\underset{w}{argmax}\,J(\boldsymbol{w})=\frac{\boldsymbol{W}^T\boldsymbol{S}_b\boldsymbol{W}}{\boldsymbol{W}^T\boldsymbol{S}_w\boldsymbol{W}} \tag{14}$$

采用拉格朗日乘子法，即可求解出最优的投影向量：$\boldsymbol{w}^*=\boldsymbol{S}_w^{-1}(\boldsymbol{\mu}_0-\boldsymbol{\mu}_1)$。在实际的运动想象模式识别中，本文将学习到的跨域特征投影至最优方向，并采用 LDA 方法进行模式识别。

3.3　运动想象 EEG 信号跨域特征学习方法流程

本文提出的基于 MI-BCI 的 EEG 信号跨域特征学习方法流程图如图 1 所示。该方法步骤如下：

(1) 根据不同个体域的 EEG 样本分布情况，采用 trial-and-error 方法选择最优的协方差距离度量方法；随后，将不同个体域的 EEG 信号样本对齐至单位阵，使得个体域的 EEG 信号样本分布更相似；

(2) 针对每个域对齐后的 EEG 信号样本提取 CSP 特征，在此基础上将其中某个被试者的 CSP 特征作为目标域，余下所有被试者的 CSP 特征作为源域，在源域和目标域间执行领域自适应方法，使得 CSP 特征之间的 *MMD* 最小化；具体来看，从 TCA，JDA，TJM 和 BDA 方法中选择最优的领域自适应方法，从多个源域中学习到最适合目标域的跨域特征。

(3) 采用 LDA 模型用于跨域特征分类，完成运动想象 EEG 信号的模式识别。

4　实验与结果分析

4.1　实验初始化

为了验证本文提出的跨域特征学习方法的可行性与有效性，本文选择 2008 BCI Competition IV 中的两个公开的 MI-BCI 数据集[36,37]，MI 1 和 MI 2a。该两个数据集基本实验范式如图 2 所示：被试者舒服坐在椅子上，面对屏幕开始实验。t=0 时，屏幕出现十字架集中注意力，伴随短暂提示音开始本次实验；t=2 时，屏幕将随机出现指导运动想象方向的箭头，被试者根据箭头指导的方向进行运动想象任务；t=6 时，被试者结束运动想象且箭头消失；此后，短暂停顿休息后进入下一次实验。

数据集 MI 1：该数据集采集了 7 名健康被试者运动想象的 EEG 数据。采集到的数据包含 59 个通道，采样率为 100 Hz。本文实验从三个运动想象任务中提取左手和右手类别，分别标

记为 Class1 和 Class2，每个用户的每个类别有 100 个样本。

数据集 MI 2a：该数据集采集了 9 名健康被试者运动想象的 EEG 数据。采集到的数据包含 22 个通道，采样率为 250Hz。本文实验从四个运动想象任务中提取左手和右手类别，分别标记为 Class1 和 Class2，每个用户的每个类别有 72 个样本。

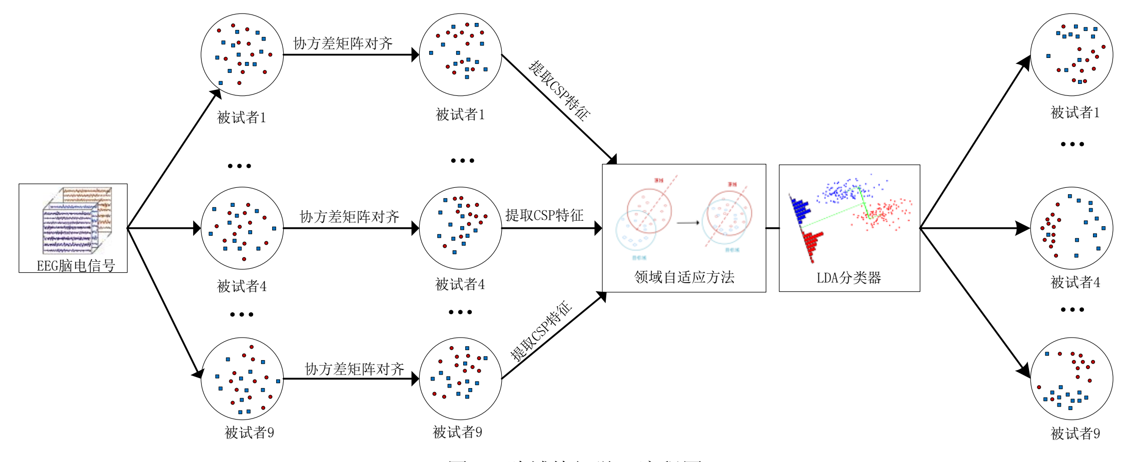


图 1 跨域特征学习流程图

Fig. 1 Cross-domain feature learning flow chart

在跨域特征学习和模式识别实验中，首先在 Matlab 上对两个数据集进行预处理，将运动想象提示箭头后的[0.5, 3.5]秒之间的采样点数作为一个运动想象样本。因此，MI 1 数据集每个样本采样时间点数为 300；MI 2a 数据集每个样本采样时间点数为 750。

实验硬件配置为 Intel(R) Core(TM) i7-9700 CPU @ 3.00GHz，8GB 内存，运行 64 位 Windows 10 系统、Matlab R2017b 以及 Python 3.7.4。协方差矩阵中心对齐方法采用代码：https://github.com/hehe91/EA，领域自适应方法代码则使用代码：https:// github.com/ jindongwang/ transfer learning / tree/master/ code。针对 JDA、BDA 和 TJM 的迭代过程中，都将迭代次数设置为 10，其他参数则针对实际数据集的测试结果设置最佳参数。本文取所有迭代结果中的最佳性能作为最终的分类结果。

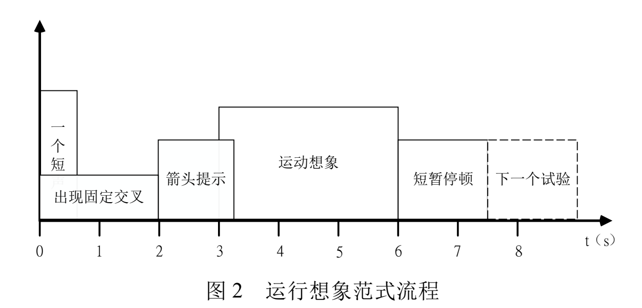


图 2 运行想象范式流程

Fig. 2 The flow chart of motor imagery paradigm

4.2 协方差矩阵中心对齐方法实验结果分析

为了验证不同均值度量计算的协方差矩阵中心对齐方法在不同数据集上的对齐效果，分别记录数据集上平均准确率最高的均值度量，以及每个被试者域最佳的均值度量，在后续实验中分别验证两种情况下的最佳跨域特征学习结果。表 1 和表 2 是分别给出了数据集 MI 1 和 MI 2a 的实验结果。其中，每行表示使用 8 种不同均值对应的同一被试者的分类准确率，每列表示不同被试者在同一均值计算的协方差矩阵中心对齐下的分类准确率，最后一行为平均结果。从表 1 和表 2 的实验结果中可以看出：

针对 MI 1 数据集，分类效果最好的是欧式均值,效果最坏的是 KL 散度均值。由于调和平均值对单个低于平均值的值特别敏感，而本质上 KL 散度均值是算数和调和均值之间的几何平均值，在 KL 散度均值下第 1 个被试者和第 4 个被试者的分类结果表现都很差，导致对整个平均效果变差。在 MI 2a 数据集下的测试结果中，分类效果最好的是 KL 散度均值，分类效果最坏的是调和均值。与 MI 1 数据集的分析过程相同，第 2 个被试者在调和均值下的分类准确率仅有 50.69%，是 8 种均值下表现最差的，因此导致最后的平均结果较差。在 KL 散度均值下表现最差的分类准确率是 54.86%，比其他均值下的最差表现都要好，单个最差值对 KL 散度均值的影响最小，因此 KL 散度均值的效果最好。

表 1 MI 1 数据集的不同均值下的分类准确率

Tab. 1 Classification accuracy under different mean values of MI1 dataset

	mean	riemann	logeuclid	logdet	wasserstein	ale	harmonic	kullback_sym
1	77.5	71.5	68	71.5	73.5	75.5	64	67
2	77.5	74.5	76.5	73.5	80.5	74	71.5	73
3	75.5	71.5	70.5	74	70	73	71	70
4	75	77.5	80	81.5	82.5	84	81	68
5	92.5	90	90.5	92	91.5	89	91.5	85.5
6	75.5	77	76.5	77	73.5	77.5	75	74.5
7	85	86.5	84.5	86	85	82.5	85	86.5
Acc	79.79	78.36	78.07	79.36	79.5	79.36	77	74.93

表 2 MI 2a 数据集的不同均值下的分类准确率

Tab. 2 Classification accuracy under different mean values of MI 2a dataset

	mean	riemann	logeuclid	logdet	wasserstein	ale	harmonic	kullback_sym
1	87.5	84.72	84.03	84.03	84.72	83.33	81.94	85.42
2	56.25	52.78	53.47	54.17	54.17	53.47	50.69	58.33
3	98.61	97.92	97.92	97.92	97.92	97.92	96.53	97.92
4	73.61	69.44	70.14	69.44	70.83	70.14	70.14	72.22
5	50	54.17	54.86	56.25	51.39	54.86	56.94	54.86
6	64.58	65.97	66.67	65.97	65.97	65.28	65.97	65.28
7	68.75	69.44	69.44	70.14	68.75	68.75	70.14	68.06
8	89.58	90.97	90.28	90.28	90.97	91.67	91.67	89.58
9	72.92	75.69	75.69	75.69	76.39	77.78	76.39	75.69
Acc	73.53	73.46	73.61	73.77	73.46	73.69	73.38	74.15

4.3 跨域特征学习实验结果分析

针对不同域对齐后的 EEG 样本，分别提取 CSP 特征，并使用 LDA 获得分类识别准确率。由于样本数量较少，分别对每一个数据集选用交叉验证法，即采用留一法依次选出目标域样本和源域样本，每轮实验依次选择一个被试者的样本集为目标域，余下的样本集为源域，每一个被试者都会被选择为目标域，最终使用分类的平均正确率作为实验结果的度量。

4.3.1 实验结果

在上文的协方差矩阵中心对齐方法实验结果基础上，本

文分别在两个数据集上继续测试样本对齐对领域自适应特征学习分类性能的提升效果。本节单独对四种领域自适应方法进行实验，以及在样本对齐后提取的 CSP 特征基础上，分别进行 TCA、BDA、JDA、TJM 四种领域自适应方法和协方差矩阵中心对齐的对比实验。为了更加明显的分析在协方差矩阵中心对齐方法下四种领域自适应方法的优劣，所有对比均在欧式均值对齐获取到的 CSP 特征上进行，表 3 和表 4 分别给出了两个数据集在不同特征学习方法上的实验结果。从表 3 和表 4 中的结果可以看出，在平均分类准确率方面:

(1) 在四种领域自适应特征学习方法的对比下，跨域特征学习的准确率高于领域自适应特征学习的准确率，实验结果充分说明协方差矩阵中心对齐方法对领域自适应特征学习方法有着至关重要的作用以及对降低个体分布差异的有效性。

(2) 不管跨域特征学习方法选择哪种核函数，本文可以发现在 MI 1 的数据集上 TJM 的方法最优，在 MI 2a 的数据集上 JDA 的方法最优。所以在欧式距离样本对齐基础上，TJM 方法更适合 MI1 数据集的样本分布，JDA 方法更适合 MI2a 数据集的样本分布。

(3) 与缺少样本对齐的领域自适应特征学习方法相比，在协方差矩阵中心对齐方法的预处理条件下，四种领域自适应特征学习方法均得到了提升。其中，在 MI 1 数据集上采用 EA+TJM 方法提高了 14.35 %，在 MI2a 数据集上采用 EA+JDA 方法提高了 3.09 %。

表 3　MI 1 数据集在不同方法下的平均准确率

Tab. 3　Average accuracy of MI 1 dataset at different methods

Kernel	样本对齐(%)	领域自适应(%)				跨域特征学习(%)			
	EA	TCA	BDA	JDA	TJM	EA+TCA	EA+BDA	EA+JDA	EA+TJM
RBF	79.79	66.29	66.14	67.5	68.86	81.21	81.00	82.43	**83.21**
Linear		66.5	65.21	67.71	67.93	81.21	80.35	82.28	82.93
Primal		66.36	66.71	67.36	68.21	81.57	79.64	83.00	83.14
Sam		66.71	66.71	67.43	68.79	81.07	81.21	82.5	82.79

表 4　MI 2a 数据集在不同方法下的平均准确率

Tab. 4　Average accuracy of MI 2a dataset at different methods

Kernel	样本对齐(%)	领域自适应(%)				跨域特征学习(%)			
	EA	TCA	BDA	JDA	TJM	EA+TCA	EA+BDA	EA+JDA	EA+TJM
RBF	73.53	72.00	72.92	73.61	74.07	73.99	74.46	**76.70**	76.47
Linear		72.53	73.00	73.61	73.69	74.69	74.86	76.16	75.38
Primal		72.22	73.00	73.23	73.69	74.31	74.46	75.62	76.00
Sam		72.00	72.92	74.15	74.31	74.38	74.77	76.08	76.62

从两个数据集总体跨域特征学习的结果可以看出，本文提出方法在每个数据集上均取得了比单一跨域特征学习角度更好的运动想象模式识别性能。为了探索单个个体域特征学习的最佳方法，本文针对每个个体域选择最优的样本对齐度量方法和 CSP 特征跨域特征方法，通过实验获得个体最优的运动想象模式识别结果。表 5 和表 6 分别给出了两个数据集个体域最优运动想象识别性能的方法。

表 5　MI 1 数据集每个个体域最优跨域特学习准确率及方法

Tab. 5　Optimal accuracy of each subject in MI 1 dataset for the cross-subject feature learning method

subjects	均值对齐	准确率/%	特征域迁移	准确率/%	最优参数	运行时间
1	mean	77.5	BDA	79	Dim=5,lambda=0.7	3.76s
2	wasserstein	80.5	JDA/TJM	84.5	Dim=2,lambda=0.7	3.60s
3	mean	75.5	TJM	79	Dim=1,lambda=0.5	3.63s
4	ale	84	BDA/JDA	86.5	Dim=6,lambda=0.4	3.73s
5	mean	92.5	TJM	95	Dim=1,lambda=0.5	3.81s
6	ale	77.5	BDA	78.5	Dim=6,lambda=0.4	3.60s
7	kullback_sym/ riemann	86.5	TJM	88.5	Dim=2,lambda=0.1	3.51s

表 6　MI 2a 数据集每个个体域最优跨域特学习准确率及方法

Tab. 6　Optimal accuracy of each subject in MI 2a dataset for the cross-subject feature learning method

subjects	均值对齐	准确率/%	特征域迁移	准确率/%	最优参数	运行时间
1	mean	87.5	JDA//TJM	92.36	Dim=1,lambda=0.5	3.0s
2	kullback_sym	58.33	JDA	60.42	Dim=4,lambda=0.1	2.96s
3	mean	98.61	JDA	99.31	Dim=1,lambda=0.4	2.93s
4	mean	73.61	JDA/TJM	77.78	Dim=2,lambda=0.1	2.78s
5	harmonic	56.94	JDA	59.03	Dim=3,lambda=0.9	3.1s
6	logeuclid	66.67	TJM	72.22	Dim=3,lambda=0.1	2.74s
7	loget/harmonic	70.14	TJM	70.83	Dim=1,lambda=0.2	2.84s
8	ale/harmonic	91.67	JDA/TJM	95.14	Dim=3,lambda=0.2	2.75s
9	ale	77.78	JDA	84.72	Dim=2,lambda=0.1	2.89s

从表 5 和表 6 中的结果中可以看出：

(1) 经过最优度量方式对齐后的样本，其 CSP 特征能够在最优域自适应方法和参数下获得再次提升，平均提升幅度为 2.43%和 3.40%；

(2) 针对不同数据集的不同个体域，样本对齐的度量方式对最终识别准确率有较大影响，花费一定时间代价获取最优的识别准确率能够极大的提升模式识别性能；

(3) 实验结果充分说明，本文提出的方法对个体分类性能的提升，面向个体域数据分布差异灵活选择合适的方法，随着算法的发展在实际应用可保证高性能和鲁棒性。

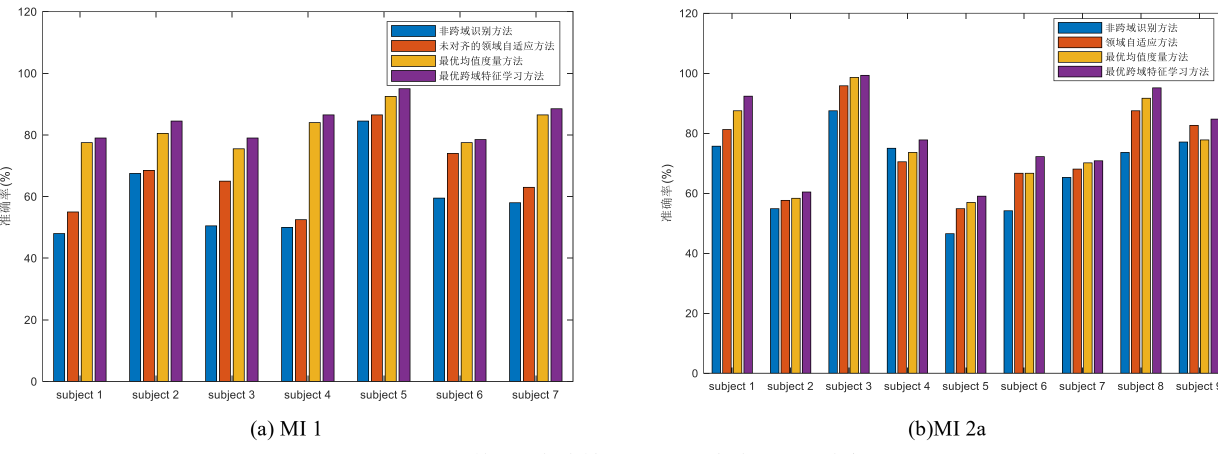
为了对比分析本文提出算法在模式识别性能上的提升，本文针对两个数据集中每个被试者对比如下种情况的运动想象识别准确率：(1)非跨域的识别方法(CSP+LDA)；(2)不经过样本对齐而直接使用领域自适应方法(CSP+DA+LDA)；(3)经过最优均值度量样本对齐后的识别方法(EA+CSP+LDA)；(4)本文提出的最优跨域特学习方法(EA+CSP+DA+LDA)。图 3 分别给出了两个数据集在上述四种情况下的运动想象识别准确率对比。从图 3 中的结果可以看出，针对每个个体域的最优跨域特征学习方法的识别准确率均高于对比的三个方法。此外，最优度量样本对齐效果基本由于经典域自适应方法。对比非跨域识别方法和本文提出的最优跨域特征学习方法，运动想象识别准确率得到了明显提升。例如 MI 1 数据集的 subject4，最优跨域特征学习方法比非跨域识别提升了 36.5%。

同时，如图 4 所示，本文还分析了在样本对齐下参数 dim 对四种域自适应特征学习方法的影响。从图 4 的结果中可以看出，参数 dim 的选择对实验影响较大。当 dim 较小时，所能承载的可迁移 CSP 特征较少，因此造成了识别准确率的降低。随着 dim 的提升，总体识别准确率先上升后下降，在 dim=4 时获得了最佳的分类性能。

随后，本文联合跨域特征学习的实验，选择 rbf 核函数，验证在它们共同作用下对分类性能的影响，MI 1 和 MI 2a 数据集上的实验结果分别如图 5 所示。从图中本文可以看出：

(1)从分类整体情况可以看出，MI 1 数据集使用欧式均值作为距离度量的领域自适应方法效果最佳，使用 KL 散度均值作为距离度量的效果最差。MI 2a 数据集使用对数欧式均值作为距离度量的效果比较好，对比协方差矩阵中心对齐方法获得的分类准确率都有了一定的提升，使用调和均值的效果整体比较差。

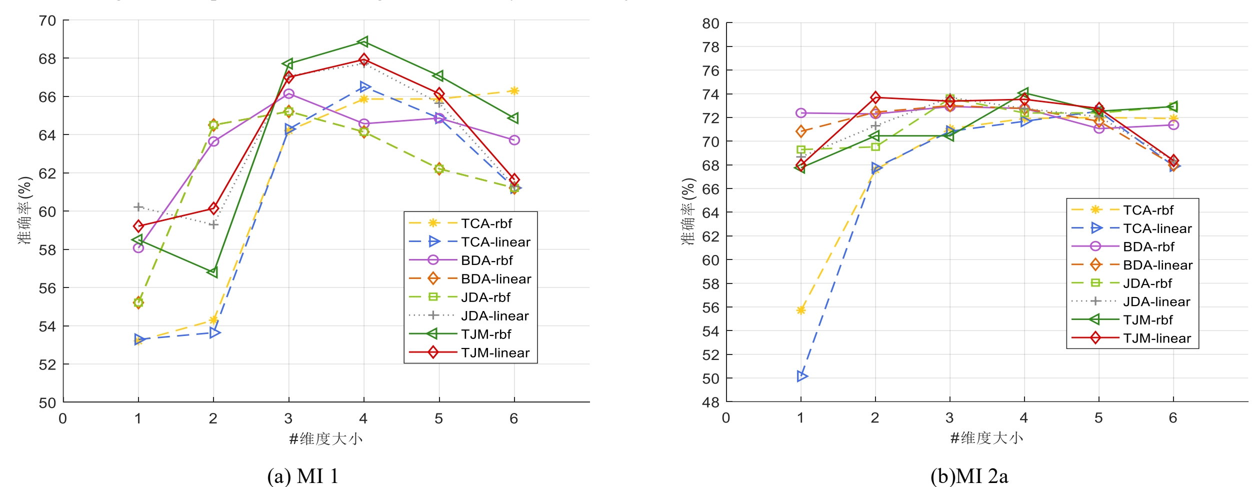
(2)MI 1 数据集中经过样本对齐的四种领域自适应方法中效果最好的是 TJM，MI 2a 数据集中效果最好的是 JDA。由此可看出，采用欧式均值和 TJM 领域自适应的特征学习方法更适合 MI 1 的样本分布，采用对数欧式均值和 JDA 域自适应的特征学习方法更适用于 MI 2a 数据集的样本分布。



(a) MI 1 (b)MI 2a

图 3 四种不同情况下每个被试者的运动想象识别准确率对比

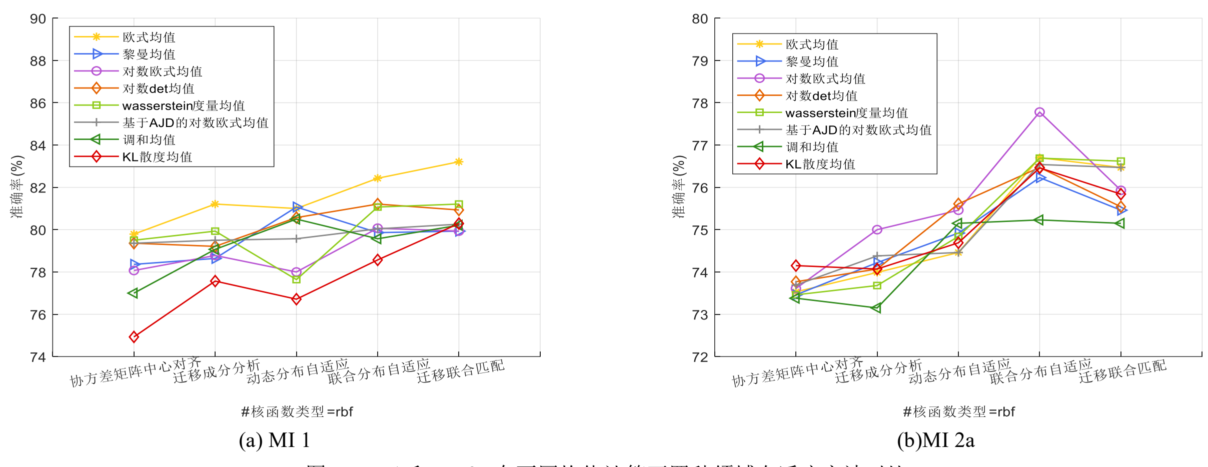
Fig. 3 Comparison of MI recognition accuracy of each subject under four different situations on MI 1 and MI 2a datasets



(a) MI 1 (b)MI 2a

图 4 MI 1 和 MI 2a 数据集在不同 dim 值下四种领域自适应方法准确率对比

Fig. 4 Comparison of accuracy of four domain adaptive methods in MI 1 and MI 2a data sets under different dim values



(a) MI 1 (b)MI 2a

图 5 MI1 和 MI 2a 在不同均值计算下四种领域自适应方法对比

Fig. 5 Comparison of four domain adaptive methods for MI1 and MI 2a dataset under different mean calculations

在实验采用的四种领域自适应方法中，TCA 方法为无监督方式且不需要多次迭代，而 BDA、JDA 和 TJM 三种算法均 需迭代获得伪标签以及求解系数矩阵。因此，图 6 给出了不同迭代次数 BDA、JDA 和 TJM 三种方法在不同核函数下的迭代准确率对比。从图中的结果可以看出，算法的效果随着迭代的次数变得更好。此外，虽然迭代对比实验中将最大迭代次数设置为 10 次，实际上绝大多数情况仅需 5 次迭代即可获得与最终迭代下匹配的平均准确率。在实际脑机接口应用场景，本文采用提出的跨域特征学习方法，仅需少量的迭代次数即可获得较好的识别准确率。

4.3.2 特征可视化

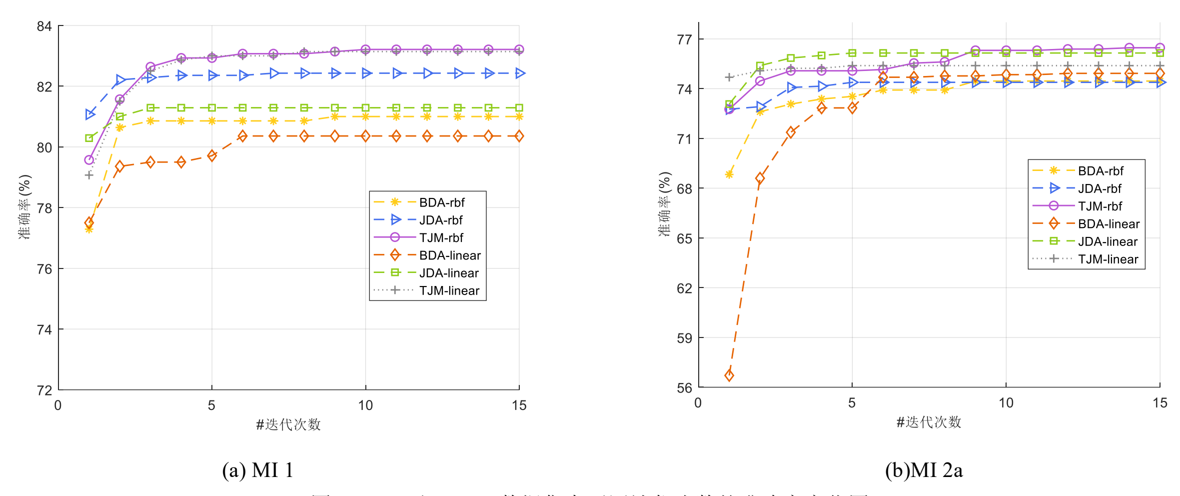
为了更直观的观察跨域特征学习对 EEG 样本和特征的变化，本文对四种领域自适应特征学习方法做了可视化处理。本文选择展示 MI 1 数据集中第 4 个被试者以及 MI 2a 数据集中第 8 个被试者的 EEG 信号跨域特征学习方法的可视化结果。可视化采用 t-SNE，将特征投影至 2 维空间并展示，图 7 和图 8 给出了数据可视化结果，图中的红色和蓝色圆点分别代表两个不同类别。

从图 7 中可以看出，进行样本对齐之后，红色圆点和蓝色三角相互之间交错比较多，说明分类的效果不是很好，但采用经过样本对齐的领域自适应方法后，红色圆点和蓝色三角的间隔变小了，交错也少了，能比较明显的区分出样本的两个类别。

从图 8 本文可以看出，在样本对齐之前，红色圆点和蓝色圆点也有比较好的分类效果，但经过样本对齐的预处理后，

图中红色圆点和蓝色三角的交错点明显少了，在采用跨域特征学习方法后，红色圆点和蓝色蓝点的区分度也有了一定程度的提升。

从上面的结果可以看出：协方差矩阵中心对齐方法可将不同类别的数据分布的边缘分布差异减小，保证分类器不受不同被试者 EEG 信号样本的边缘分布影响。此外，TCA 方法可在协方差矩阵中心对齐方法基础上，将相同被试者不同类别信号的边缘分布差异减小，而 JDA 和 BDA 方法则不但减小边缘分布，更减小了标签参与的条件分布差异，因此能够更好的提升平均识别准确率。TJM 方法保证了特征选择的稀疏性，因此在可视化的结果上，可以被选择后的特征具有稀疏性，且让不同类别的特征分离的差异更大。



(a) MI 1 (b)MI 2a

图 6 MI1 和 MI 2a 数据集中不同迭代次数的准确率变化图

Fig.6The diagram of accuracy variation under different iterations

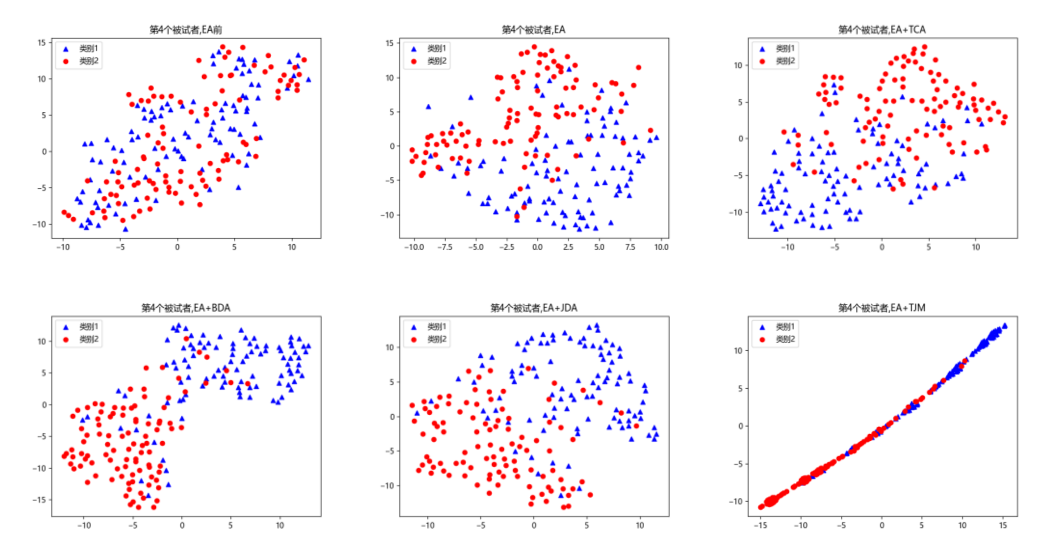


图 7 MI 1 数据集中第 4 个被试者经过 T-SNE 的可视化结果

Fig.7MI 1 dataset represents the visualization results of the fourth subject through T-SNE

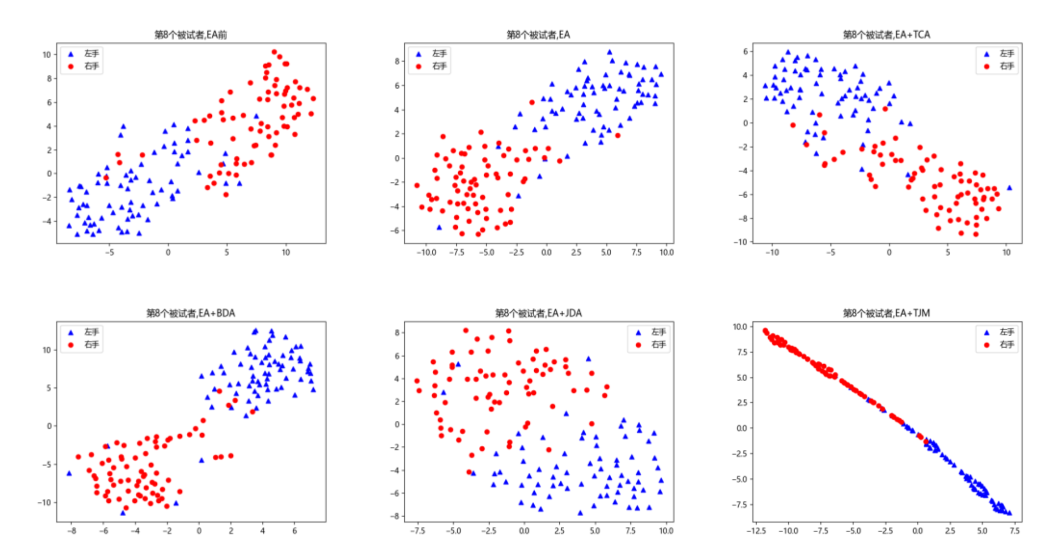
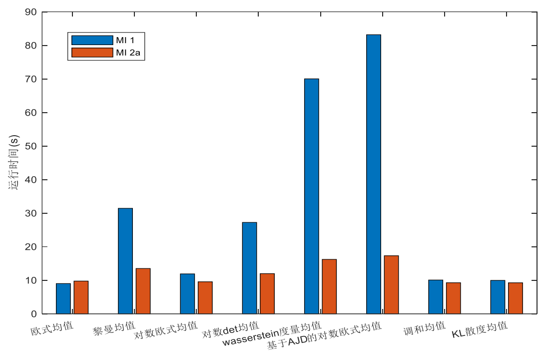


图 8 MI 2a 数据集中第 8 个被试者经过 T-SNE 的可视化结果

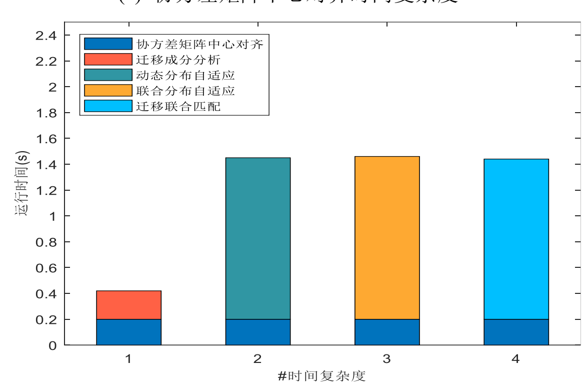
Fig.8MI 2a dataset represents the visualization results of the fourth subject through T-SNE

4.4　时间复杂度分析

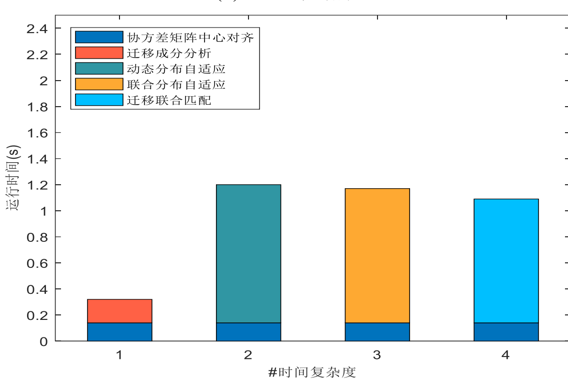
图 9 给出了本文方法的运行时间复杂度对比结果。其中，图 9(a)展示了不同均值度量平均运行时间。从该柱状图的结果可以观察到，两个 MI 数据集中在欧式均值度量下的运行时间最短，而基于 AJD 的对数欧式均值运行时间最长，前者快 2~10 倍。此外，MI 1 数据集数据量更大，运行时间更久。



(a) 协方差矩阵中心对齐时间复杂度



(b) MI 1 数据集



(c)MI 2a 数据集

图 9 八种均值和四种域自适应方法的运行时间对比

Fig.9Running time comparison of eight mean and four domain adaptive methods

以欧式均值度量为基础，图 9 (b)(c)给出了不同域自适应方法运行时间。在 BDA、JDA 和 TJM 三种算法中迭代次数为 5 次时已经接近最优解，从图 8(b)(c)中结果可以看出，TCA 的运行时间最短，相对而言 BDA、JDA 和 TJM 的运行时间相差不大，且均在 1.5 秒以内。因此，所提出对脑电信号识别方法的响应时间足以应对运动想象刺激范式的流程，可应用于在线脑机接口。上述结果表明，最优跨域特征学习方法的平均时间复杂度在可接受范围内，同时在实际的 MI-BCI 应用中使用跨域特征学习方法时，需根据实际情况同时考虑准确率和时间复杂度，选择适合场景的协方差中心均值度量及领域自适应方法。

目前，现有脑机接口系统(如外骨骼辅助康复)遭遇新被试者时，往往采集大量样本训练模式识别模型。从图 2 可看出，单个样本至少需要 4 秒时间，且通常需要 200-300 个样本才能稳定识别。采用本文提出的跨域特征学习方法，新的被试者仅需采集少量样本(100 个以内)，即可通过跨域特征学习使用已有使用者的脑电信号样本集，且随着使用者增多，可使用的脑电信号样本集也越多，对模式识别的效果也越好。此外，新的被试者所采集少量样本的最优选择，即可使用现有样本集获得更高的模式识别性能，其时间代价在可接受范围内。

5　结束语

针对运动想象个体域 EEG 信号获取代价高且特征分布差异大的特性，本文提出了一种运动想象脑电信号的跨域特征学习方法，用于提升跨个体域脑电信号模式识别性能。该方法首先选择最优的协方差距离度量方法，对齐各域样本集并提取 CSP 特征。其次，在 CSP 特征基础上，选择合适的领域自适应方法，从多源域中学习出最适合目标域的跨域特征，最后采用 LDA 分类完成模式识别任务。实验结果表明，所提出的方法明显可提升模式识别准确率。跨域特征学习能够极大的提升 BCI 应用中的样本数量和质量，提升控制外部设备的性能，然而现有方法存在一个缺点：采用离线方式进行。在实际应用中，源域样本或许能够长期积攒，但是目标域样本不断产生，无法直接计算协方差矩阵中心以及 CSP 特征。今后的工作集中在：探索迭代式协方差计算方法，设计增量迭代式的跨域特征学习方法，并将迭代式跨域特征学习应用至在线 MI-BCI 中。

参考文献：

[1] Lance B J, Kerick S E, Ries A J, *et al.* Brain-computer interface technologies in the coming decades [J]. Proceedings of The IEEE, 2012, 100 (1): 1585-1599.

[2] Wolpaw J R, Birbaumer N, Mcfarland D J, *et al.* Brain-computer Interfaces for Communication and Control [J]. Clinical Neurophysiology, 2002, 113 (6): 767-791.

[3] 周鹏. 基于运动想象的脑机接口的研究 [D]. 天津: 天津大学, 2007.

[4] Pfurtscheller G, Müeller-Putz, Gernot R, *et al.* Rehabilitation with Brain-Computer Interface Systems [J]. COMPUTER, 2008, 41 (10): 58-65.

[5] Hashimoto Y, Kakui T, Ushiba J, *et al.* Development of Rehabilitation System with Brain-Computer Interface for Subacute Stroke Patients [C]// Proc of the IEEE International Conference on Systems Man and Cybernetics Conference Proceedings. Miyazaki: IEEE Press, 2018: 51-56.

[6] Kasashima-Shindo Y, Fujiwara T, Ushiba J, *et al.* Brain-Computer Interface Training Combined with Transcranial Direct Current Stimulation in Patients with Chronic Severe Hemiparesis: Proof of Concept Study [J]. Journal of Rehabilitation Medicine, 2015, 47 (4): 318-324.

[7] Blankertz B, Tomioka R, Lemm S, *et al.* Optimizing Spatial Filters for Robust EEG Single-Trial Analysis [J]. IEEE Signal Processing Magazine, 2008, 25 (1): 41-56.

[8] He He, Wu Dongrui. Transfer Learning Enhanced Common Spatial Pattern Filtering for Brain Computer Interfaces (BCIS): Overview and A New Approach [C]// Proc of the 24th International Conference on Neural Information Processing. Guangzhou: ICONIP Press, 2017: 811-821.

[9] 庄福振, 罗平, 何清, 等. 迁移学习研究进展 [J]. 软件学报, 2015, 26 (1): 26-39. (Zhuang Fuzhen, Luo Ping, He Qing, *et al.* Survey on Transfer LearningResearch [J]. Journal of Software, 2015, 26 (1): 26-39.)

[10] Pan S J, Tsang I W, Kwok JT, *et al.* Domain Adaptation via Transfer Component Analysis [C]// Proc of the 21st International Joint Conference on Artificial Intelligence. Pasadena: IJCAI Press, 2009: 1187-1192.

[11] Long Mingsheng, Wang Jianmin, Ding Guiguang, *et al.* Transfer Feature Learning with Joint Distribution Adaptation [C]// Proc of the IEEE International Conference on Computer Vision. Sydney: IEEE Press, 2013: 2200-2207.

[12] Long Mingsheng, Wang Jianmin, Ding Guiguang, *et al.* Transfer Joint Matching for Unsupervised Domain Adaptation [C]// Proc of the 27th IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE Press, 2014: 1410-1417.

[13] Wang Jindong, Chen Yiqiang, Hao Shuji, *et al.* Balanced Distribution Adaptation for Transfer Learning [C]// Proc of the 17th IEEE International Conference on Data Mining. New Orleans: IEEE Press, 2017: 1129-1134.

[14] 李晶晶, 孟利超, 张可, 等. 领域自适应研究综述 [J]. 计算机工程, 2021, 47 (06): 1-13. (Li Jingjing, Meng Lichao, Zhang Ke, *et al.* A Review of Domain Adaptation Research [J]. Computer Engineering, 2021, 47 (06): 1-13.)

[15] 田磊, 唐永强, 张文生. 两阶段领域自适应学习 [J]. 模式识别与人工智能, 2019, 32 (09): 773-784. (Tian Lei, Tang Longqiang, Zhang Wensheng. Two-Stage Domain Adaptive Learning [J]. Pattern Recognition and Artificial Intelligence, 2019, 32 (09): 773-784.)

[16] Dai Mengxi, Zheng Dezhi, Liu Shucong, *et al.* Transfer Kernel Common Spatial Patterns for Motor Imagery Brain-Computer Interface Classification [J/OL]. Computational and Mathematical Methods in Medicine, 2018, 2018 (9): Article ID 9871603. (2018-04-17) [2022-01-15]. https://doi. org/10. 1155/2018/9871603.

[17] Albalawi H, Song Xiaomu. A Study of Kernel Csp-based Motor Imagery Brain Computer Interface Classification [C]// Proc of the IEEE SignalProcessing in Medicine and Biology Symposium, New York: IEEE Press, 2012: 1–4.

[18] Long Mingsheng, Wang Jianmin, Sun Jiaguang, *et al.* Domain Invariant Transfer Kernel Learning [J]. IEEE Trans on Knowledge and Data Engineering, 2015, 27 (6): 1519-1532.

[19] Jayaram V, Alamgir m, Altun Y, *et al.* Transfer Learning in Brain-Computer Interfaces [J]. IEEE Computational Intelligence Magazine, 2015, 11 (1): 20-31.

[20] Azab A M, Mihaylova L, Ang K K, *et al.* Weighted Transfer Learning for Improving Motor Imagery-Based Brain-Computer Interface [J]. IEEE Trans on Neural Systems and Rehabilitation Engineering, 2019, 27 (7): 1352-1359.

[21] Hossain I, Khosravi A, Hettiarachchi I, *et al.* Multiclass Informative Instance Transfer Learning Framework for Motor Imagery-Based Brain-Computer Interface [J]. Computational Intelligence and Neuroscience, 2018, 2018 (9): Article ID 6323414. (2018-02-22) [2022-01-15]. https://doi. org/10. 1155/2018/6323414.

[22] Zanini P, Congedo M, Jutten C, *et al.* Transfer Learning: A Riemannian Geometry Framework with Applications to Brain-Computer Interfaces [J]. IEEE Trans Biomed Eng, 2017, 65 (5): 1107-1116.

[23] Barachant A, bonnet S, Congedo M, *et al.* Multiclass brain-computer interface classification by Riemannian geometry [J]. IEEE Trans on Biomedical Engineering, 2012, 59 (4): 920–928.

[24] Yair O, Dietrich F, Talmon R, *et al.* Domain Adaptation with Optimal Transport on the Manifold of SPD matrices [J]. IEEE Trans on Signal Processing, 2019, 67 (7): 1797–1811.

[25] He He, Wu Dongrui. Transfer Learning for Brain-Computer Interfaces: A Euclidean Space Data Alignment Approach [J]. IEEE Trans on Biomedical Engineering, 2020, 67 (2): 399-410.

[26] Rodrigues P L C, Jutten C, Congedo M. Riemannian Procrustes Analysis: Transfer Learning for Brain-Computer Interfaces [J]. IEEE Trans on Bio Medical Engineering, 2018, 66 (8): 2390-2401.

[27] Zhang Wen, Wu Dongrui. Manifold Embedded Knowledge Transfer for Brain-Computer Interfaces [J]. IEEE Trans on Neural Systems and Rehabilitation Engineering, 2020, 28 (5): 1117-1127.

[28] Singh A, Lal S, Guesgen H W. Small Sample Motor Imagery Classification Using Regularized Riemannian Features [J]. IEEE Access, 2019, 7 (1): 46858-46869.

[29] Wu Dongrui, Xu Yifan, Lu Bao-liang. Transfer Learning for EEG-Based Brain-Computer Interfaces: A Review of Progress Made Since 2016 [J]. IEEE Trans on Cognitive and Developmental Systems, 2020, 2020 (16): Article ID 9134411. (2020-07-07) [2022-01-15]. https://ieeexplore. ieee. org/document/9134411.

[30] Wu Hao, Niu Yi, Li Fu, *et al.* A Parallel Multiscale Filter Bank Convolutional Neural Networks for Motor Imagery EEG Classification [J]. Frontiers in Neuroscience, 2019, 13 (1): 1-9.

[31] Wang Peitao, Lu Jun, Zhang Bin, *et al.* A Review on Transfer Learning for Brain-Computer Interface Classification [C]// Proc of the 5th International Conference on Information Science and Technology. Changsha: ICIST Press, 2015: 315-322.

[32] Ren Jiangtao, Liang Zhou, Hu Shaofeng. Multiple Kernel Learning Improved by MMD [C]// Proc of the 6th International Conference on Advanced Data Mining and Applications. Chongqing: ADMA Press, 2010: 63-74.

[33] 赵鹏, 吴国琴, 刘慧婷, 等. 基于特征联合概率分布和实例的迁移学习算法 [J]. 模式识别与人工智能, 2016, 29 (08): 717-724. (Zhao Peng, Wu Guoqin, Liu Huiting. Transfer Learning Algorithm Based on Feature Joint Probability Distribution and Instance [J]. Pattern Recognition and Artificial Intelligence, 2016, 29 (08): 717-724.)

[34] Mishuhina V, Jiang Xudong. Feature Weighting andRegularization of Common Spatial Patterns in EEG-Based Motor Imagery BCI [J]. IEEE Signal Processing Letters, 2018, 25 (6): 783-787.

[35] Subasi A, Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machines [J]. Expert Systems with Applications, 2010, 37 (12): 8659-8666.

[36] Tangermann M, Muller KR, Aertsen A, *et al.* Review of the BCI Competition IV [J]. Frontiers of Neuroscience, 2012, 6 (1) . (2012-07-13) [2022-01-15]. https://www. frontiersin. org/articles/10. 3389/fnins. 2012. 00055/full

[37] Ang K K, Chin Z Y, Wang C C, et al. Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b[J]. Frontiers of Neuroscience,2012,6(1).(2012-03-29)[2022-01-15]. https: // www . frontiersin.org/ articles/10.3389/ fnins.2012.000 39